

Streaming Panoramic Video

Related Application:

The present application is a continuation in part of application 60/210,374 filed 6/9/2000.

Field of the Invention:

The present invention relates to the transmission of panoramic images and more particularly to transferring portions of panoramic images from a server to a client using "video streaming".

Background of the Invention:

It is well known that special provisions are required when viewing panoramic images on a computer display. If an entire panoramic image is projected on a computer display, the image is necessarily distorted. Panoramic images are generally viewed using a viewer program which renders (i.e. displays) a portion of the panorama on a screen or display. The portion of the panorama that is displayed is generally termed a "view window". Generally viewer programs provide a mechanism (such as a mouse) that can be used to select the desired portion of the panorama frame that constitutes the view window.

A panoramic video (or a panoramic movie) is a series of panoramic frames, each of which contains a panoramic image. Co-pending patent application 09/310,715 filed 05/12/99 entitled "Panoramic Movies which Simulate Movement Through Multidimensional Space" describes a system for displaying a panoramic video by displaying a view window that displays in sequence substantially the same view window from a series of panoramic images. The view window only gradually changes location between frames as the viewer chooses to change the direction of view.

Storing a panoramic video requires a great deal of storage, hence, a large amount of bandwidth is required in order to stream panoramic images from a web server to a client. The present invention is directed to reducing the bandwidth required to stream panoramic images from a web server to a client. With the present invention panoramic images can be

1 streamed from a web server to a client over a lower bandwidth connection or with greater
2 image quality, size, and/or frame rate.

3
4 The MPEG video compression standard provides a "slicing" mechanism. This mechanism
5 is generally used in order to facilitate error correction. The present invention utilizes the
6 slicing mechanism in the MPEG video compression standard to reduce the bandwidth
7 required to stream panoramic video from a web server to a client.

8 9 **Summary of the Invention.**

10 The present invention streams panoramic images from a server to a client. The system
11 utilizes a special module at the client and a special module in the server. The special
12 modules may be plug-ins for commercially available streaming programs. The special
13 module at the client provides the functions that are typically provided by a conventional
14 panorama viewer program and it also communicates with the module at the server to
15 specify which portion of the panorama should be streamed to the client. The special
16 module at the client has the ability to accept data that represents a portion of a series of
17 panoramic frames, to decompress the data, to select the data that constitutes an
18 appropriate view window and to render (i.e. display) a portion of each frame on a screen or
19 display.

20
21 The server module selects particular slices that constitute a region of interest in the
22 panorama and these slices are sent to the client. At the client, the user may select
23 navigation commands such as pan left, pan right, pan up, pan down, roll left, roll right, zoom
24 in, zoom out or a combination of these or other commands to change the view window.
25 When the location of the view window is changed by more than a threshold amount, the
26 client sends a command back to the web server. In response to the commands from the
27 client, the server module adjusts the selection of the slices that are streamed from the
28 server to the client. There may be many clients receiving information from a particular
29 server and for every client, the module at the server maintains session information and
30 streams appropriate information to that client.

31 32 **Brief Description of the Figures:**

1 Figure 1 is a block diagram of first embodiment of the invention.

2 Figures 2A to 2D illustrate the movement of a region of interest and a view window in a
3 panoramic image.

4 Figure 3 is a block diagram of the program in the browser plug in.

5 Figure 4 is a block diagram of the program in the server plug in.

6 Figure 5 illustrates the shape of a view window relative to the slices in a panorama.

7 Figure 6 shows an alternate form of panoramic image.

8 Figure 7 shows an alternate embodiment of the invention wherein two different streams are
9 being transmitted from the server to different clients.

10 Figure 8 shows another alternate embodiment of the invention which utilizes a different type
11 of server.

12 Figure 9 shows an embodiment of the invention where the entire invention is operating on a
13 single computer.

14
15 **Detailed Description:**

16 A first preferred embodiment of the invention is shown in Figure 1. In this embodiment
17 panoramic images are streamed from a server 100 to a client 150 over a network 120. The
18 network 120 could for example be the Internet. While only a single client 150 is shown it
19 will be understood by those skilled in the art that a single server 100 could provide data to a
20 large number of clients 150.

21
22 The data streamed from server 100 to client 150 could for example be data from a
23 panoramic movie of the type shown in co-pending application 09/310,715 filed 05/12/99
24 entitled "Panoramic Movies which Simulate Movement Through Multidimensional Space",
25 the content of which is hereby incorporated by reference. A panoramic movie consists of a
26 series of panoramic images. Such a series of panoramic images could for example be a
27 series of panoramas recorded by a multi lens camera which is moving along a street. A
28 panorama is normally displayed by allowing a user to select a view window (i.e. the
29 direction in which the user is looking). In a panoramic movie, this view window can change
30 direction as a series of frames is projected. That is, with a panoramic movie, the user has
31 the option of selecting the direction of view. The location of the view window in the
32 panorama changes as the user changes direction of view.

1
2 With the present invention an entire panorama is not streamed from the server 100 to the
3 client 150. Only that portion of the panorama (called a region of interest) that includes the
4 view window and a surrounding region (i.e. a guard band) is streamed from the server 100
5 to the client 150. That is, the region of interest that is streamed from the server to the client
6 includes the view window and a guard band around the view window. The user is provided
7 with controls (e.g. a mouse 159) whereby the user can change the location of the view
8 window in the panorama, that is, the user can change the area of the panorama that is
9 being displayed. When the user changes the location of the view window by more than a
10 threshold amount, the client sends a command to the server to change the location of the
11 region of interest.

12
13 Data in the entire region of interest is transmitted from the server 100 to the client 150. The
14 client therefore has the entire region of interest immediately available for display. The
15 guard band surrounding the view window provides data that is immediately available for
16 display at the client when the user moves the view window. Thus, the user can change (to
17 some degree) the location of the view window in the panorama and the data needed to
18 provide the changed display is immediately available without having to wait for the server to
19 send different data.

20
21 Without the present invention, one could achieve the same result by streaming entire
22 panoramas from the server 100 to the client 150; however, this would require significantly
23 more bandwidth than is required by the present invention. Alternatively, only the data that
24 is in the view window could be streamed from the server to the client; however, if this were
25 done when the user gives a command to change the viewing direction (i.e. the location of
26 the view window in the panorama), the command from the user would have to go from the
27 client to the server and the server would have to begin streaming different data to the client
28 150. This would result in a delay between when the user gives a command and when the
29 view window actually changes. It is noted that this delay is exacerbated by the fact that
30 streaming systems normally buffer data at the server and at the client. Buffering is required
31 for a number of reason including the need for multiple frames in order to perform de-
32 compression.

Figures 2A to 2D illustrate how changes in the location of the view window generates changes in the area of interest that is streamed from the server to the client. Figure 2A illustrates one panorama 214. The panorama is divided into areas 214A, 214B etc. There is a region of interest 215 and the view window is 216. It is noted that the size of the areas in Figures 2A to 2D is exaggerated for purposes of illustration and they do not constitute actual MPEG slices. The actual sizes are explained later.

Figures 2A to 2D illustrate four frames in a panoramic video. It should be noted that the four frames shown in Figures 2A to 2D are not necessarily adjacent sequential frames. That is, out of a series of thirty frames, the frames (i.e. the panoramas) shown may be the first, tenth, twentieth and thirtieth frames. The changes in the intermediate frames will be a portion of the changes shown in Figures 2A to 2D.

For simplicity in illustration and ease of explanation in Figures 2A to 2D, the areas are shown as being square and the size of the view window is shown as coinciding with the size of an area. The actual size of the areas and actual shapes will be explained later. Furthermore, a panorama would normally include an image. For ease of illustration, in Figures 2A to 2D the areas are shown without showing the actual image.

The entire panorama 214 is not transmitted from the web server 100 to the client 150. Only a region of interest 215 from each frame is transmitted from the server to the browser. The region of interest 215 includes the particular view window 216 that is being displayed to the user.

When a user is looking at a particular view window in a panorama, the user might decide to change the location of the view window in the panorama. That is, the user might want to position the view window in a different part of the panorama so that a different part of the panorama will be visible on the display. The term "pan" means that a user changes the location of the view window in one direction or another.

1 Since the region of interest 215 includes a "guard band" surrounding the view window 216,
2 and since the entire region of interest 215 is transmitted to the client 150, the data is
3 available at the client 150 to allow the user to change the location of the view window (i.e.
4 to change the portion of the panorama being displayed) without the need for any
5 communication to the server 100.

6
7 Figure 2B illustrates the view window 216 moving to the right. As the user changes the
8 location of the view window 216, (i.e. as the user changes the portion of the panorama
9 being displayed) the region of interest 215 is changed as shown in Figure 2C. Motion by a
10 user generally continues in the same direction for some time so the user might arrive at the
11 location shown in Figure 2D.

12
13 Each time the user changes the location of the view window by an amount which exceeds a
14 certain threshold (which can be set depending on factors discussed later), the client 150
15 sends a message to the server 100 notifying the server of this change. When the server
16 receives a signal indicating that the location of the view window has changed, the server
17 changes (if appropriate) the particular slices being sent to the browser (i.e. the slices that
18 constitute the region of interest) so that the slices transmitted always include the view
19 window plus a guard band. Thus, the server continues sending a particular region of
20 interest from each frame until notified to change by the client. A user can pan within this
21 region of interest without waiting for the server to change the portion of the panorama that
22 is being streamed from the server to the client.

23
24 Frames in a panoramic video are generally sent at a rate of thirty frames per second. Thus,
25 the region of interest from a significant number of frames may be transmitted before the
26 server receives and reacts to a command to change the region of interest. Since the guard
27 band surrounds the view window, the user can change the location of the view window (to
28 some extent) before the server has a chance to react to a command to change the location
29 of the region of interest.

30
31 The size of the guard band does not have to be of a fixed size, or symmetrical around the
32 region of interest. The guard band may be larger in an expected or usual direction of

1 panning. For example, the guard band may be larger on the left and right sides of the view
2 window, than at the top or bottom. The size of the guard band can be adjusted to an
3 appropriate amount by tracking the history of usage by each particular user and the
4 bandwidth available. Transmitting a larger region of interest requires more bandwidth.
5 Furthermore, the viewer program may limit the rate at which the image is panned. This
6 would be done to attempt to preserve smooth panning in return for a reduced pan rate.

7
8 The panoramic frames are compressed by the server 100 using standard MPEG
9 compression. The MPEG standard specifies that slices are always 16 pixels high and
10 that the width of a slice is a multiple of 16 pixels, up to the entire width of the frame.
11 With the present invention it has been found that with a frame that is 2K by 1K, the
12 frame can be divided into 8 slices horizontally each slice being 16 pixels tall, and
13 256 pixels wide. Thus, there would be 512 slices for each frame.

14
15 It is noted that "Slicing" is a term used in the MPEG 2 standard. In the MPEG 4
16 standard, the slicing mechanism is part of the error correction and concealment
17 section of the standard, and it is known as "inserting resynchronization markers", or
18 "resynchronization mechanism". While the terms used in the two standards differ
19 somewhat the actual implementation is identical, since MPEG 4 carries over all of
20 MPEG 2's implementation. Herein the term "slice" from the MPEG 2 standard is
21 used; however, it should be understood that as used herein the term "slice" is
22 intended to refer to "slices" from the MPEG 2 standard and to the equivalent
23 mechanism in other MPEG standards.

24
25 MPEG compression uses "I" frames (Intra frames), "P" frames, and/or "B" frames. The I
26 frames contain all of the information needed to reconstruct a single image. P (Predictive)
27 frames copy the closest matching block of pixels from the preceding I or P frame, and add a
28 (hopefully small) correction to create blocks. B (Bi-directional) frames are similar to P
29 frames, but can also copy blocks from the future I or P frame, and/or can average a
30 preceding and future block to create a block in the frame being constructed. I frames are
31 relatively large, P frames are typically smaller, and B frames are usually the smallest. The

1 construction and definition of I frames, B frames, and P frames is set out in the publicly
2 available MPEG standards. The use of either B or P frames is chosen depending upon
3 whether or not reverse motion is desired.

4
5 The I frames are considerably larger than the B or P frames. Thus, in the first embodiment,
6 only slices from the region of interest in the I frames is transmitted from the client to the
7 server and the entire B or P frames are transmitted. Alternatively only slices in the region of
8 interest from the B frames could be transmitted. However, it is noted that the number of
9 slices transmitted from the I or P frames may be larger than the number of slices
10 transmitted from the B frames. The reason for this is that only the slices in the B frames
11 that are in the region of interest need be transmitted. With respect to the I and P frames,
12 both the slices in the region of interest and the slices needed by their dependent P and B
13 frames must be transmitted. This imposes a requirement that when encoding P and B
14 frames, blocks of pixels may only be copied from the corresponding slice of the referenced I
15 or P frame, and perhaps the adjacent slice as well.

16
17 When motion is stopped and a user focuses on one frame, the bandwidth can be used to
18 transmit the additional information and to store this additional information in a buffer just in
19 case it is needed. In the situation where a user stops the motion of the video, freezing the
20 view window on a portion of one frame, the system can transmit the entire panorama (or a
21 relatively large portion thereof) from the server to the browser, allowing the user full
22 freedom to pan, tilt, etc., at full speed within the current panorama without need to send
23 commands to the server. If the entire panorama (or a large portion thereof) is stored in a
24 buffer at the client machine, moving the view window can be changed over a larger region
25 more quickly.

26
27 In the first preferred embodiment of the invention shown in Figure 1, server 100 consists of
28 a conventional server platform with the "Microsoft Windows 2000" operation system 101.
29 The system includes the commercially available "Real System Server 8" program 103 which
30 is commercially available from RealNetworks Inc. The system includes a memory
31 subsystem 102 which stores panoramic videos. The overall streaming operation is handled
32 by the Real System Server 8; however, when the system is asked to stream a panoramic

1 video, the file is passed to plug-in 105. The system shown in figure 1 also includes the
2 Microsoft Internet Information Server 104. The Microsoft Internet Information Server 104 is
3 not used during the streaming operation; however, it may handle a web site that allows a
4 user to request that a particular panoramic movie be streamed. That is, a web site may list
5 a set of available panoramic movies. When a user clicks on one of the listed movies, the
6 system retrieves that files and begins sending the images to plug in 105

7
8 Figure 4 is a program block diagram showing the operation performed by plug-in 105. The
9 frames are stored in compressed format in memory system 102. When the system is asked
10 to stream a panoramic video, the panoramic frames are passed to the plug-in 105 from real
11 player 8. The system starts by transmitting a default region of interest from the panoramas
12 with the view window located at a default location. Commands to change the region of
13 interest are received from the client as indicated by block 401. As indicated by block 404,
14 the slices which form the region of interest 216 are selected. As indicated by block 405, the
15 selected slices are passed to the Real System 8 for transmission to the browser.

16
17 In the embodiment shown in Figure 1, the client 150 consists of a personal computer 151
18 with the Microsoft Windows operating system 152, the Microsoft Internet Explorer Browser
19 153, and the Real Player 8 Plus program which is commercially available for Real Networks
20 Inc. The system includes a user input device 159 such as a mouse. Finally the client 150
21 includes a plug in 155 which handles panoramic images.

22
23 Figure 3 is a block diagram of the program in plug in 155. Plug in 155 receives inputs from
24 the user and from Real Player 8 as indicated by blocks 301 and 302. As indicated by block
25 303 the slices received from the server 100 are decompressed and stored. As indicated by
26 block 304, the slices which constitute the view window are selected and this image is
27 rendered as indicated by block 305 and sent to the real player 8 port for display as
28 indicated by block 306. The view window from the panorama is rendered in a perspective
29 correct manner using the transformation known in the prior art for this purpose. Once the
30 view window is determined the selection and rendering of the appropriate data is similar to
31 the operation of many panoramic viewing programs.

The "Real System Server 8" and the "Real Player 8", that is units 103 and 154 shown in Figure 1, have what is called a "back channel". The back channel is a communication channel that is separate from the channel used to stream the video frames. The back channel can accept data from the Real Player and send it to the Real System Server, or it can accept data from the Real System Server and send it to the Real Player. The back channel is regularly used to send a command such as Stop and Reverse from the player to the server. It is this back channel that is used to send data from client 150 to server 100 to instruct the server to change the region of interest. Naturally the plug-ins 105 and 155 includes the other conventional components that are specified by documentation for the plug in specification for the Real Player 8 and the Real System Server 8.

It is noted that the size of a view window will typically be on the order of the size of about twenty to eighty MPEG slices. As is know in the art, the actual size depends upon the size of the display and the characteristics of the particular viewer software. The size of the guard band around the view window will have a size in the range of 10 to 50 MPEG slices. Thus the areas shown in Figures 2A to 2D are the size of about ten to fifty MPEG slices.

As indicated by block 307, the plug-in determines if different slices are required to constitute the appropriate area of interest 215. This is done according to the following logic where "t" "x", and "n" are variables the value of which is set as discussed below.

- a) Has the view window changed by more a threshold amount "t" ?
 - b) If the location of the view window has changed determine direction of movement.
 - c) When view window has moved by the threshold amount, move the region of interest "n" slices in that direction.
 - d) No further movement of the region of interest is necessary until the view window has moved a distance equal to "x" amount.
 - e) When the view window has moved "x" amount, revert to step "a".
 - f) If direction of movement changes, revert to step "b".
 - g) If "action stopped" and user stops on a particular frame, instruct the server to send other slices to in effect enlarge the region of interest available at the client.
- This data is stored at the client.

1 The variables "t", "x" and "n" can be initially set to default values and changed to suit the
 2 actions of a particular user and system. For example, the value of "t", "x" and "n" can be in
 3 the order of the size of 5 to 50 slices. They can be set to one size and maintained at that
 4 size throughout a session or they can be changed during a session to make the system
 5 react to existing conditions. Initially they may be set to the value which is the size of 20
 6 slices. If, for example, it is found that the system is experiencing a large amount of latency
 7 from when a command is send from the client to the server and when the server reacts, the
 8 values may be increased.

10 The above calculation takes place for both movement in the x direction and for movement
 11 in the "y" direction. As indicated by block 309 the instructions to change the slices that
 12 constitute the area of interest 215 are sent from the client 150 to the server 100.

14 As a specific example of how the system operates, consider a sequence of 500 panoramas
 15 in a panoramic move. Each panorama is 360 degrees in the horizontal direction and 180
 16 degrees in the vertical direction, represented as an image with 2,048 (2K) pixels in the
 17 horizontal direction and 1,024 (1K) pixels in the vertical direction, for a total of 2,097,152
 18 (2M) pixels per panorama.

20 When compressed this movie might consist of one "I" frame followed by nine "B" frames,
 21 followed by another "I" frame, nine "B" frames, etc. Each frame would be divided into 1024
 22 slices, 16 slices horizontally by 64 slices vertically, each slice having a size of 16 pixels
 23 vertically by 128 pixels horizontally.

25 Assume a default view window centered vertically and horizontally within the panorama of
 26 approximately 90 degrees horizontally by 45 degrees vertically. Ignoring, for simplicity, the
 27 slight panoramic distortion that occurs about the horizon of the stored panoramic image, the
 28 view window would be represented by a region of 512 ($2048/(360 \text{ degrees}/90 \text{ degrees})$)
 29 pixels horizontally by 256 ($1024/(180 \text{ degrees}/45 \text{ degrees})$) pixels vertically, or 4 slices
 30 horizontally by 16 slices vertically. Assuming a guard band of one slice all the way around
 31 the view window, the initial region of interest of each frame having a size of 6 ($4+2$) slices
 32 by 18 ($16+2$) slices would be transferred from the server to the client.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32

In a simple example, if the user moved the view window 45 degrees to the right, the client would tell the server to shift the region of interest by two columns of slices to the right. If the user moved the window only 10 degrees to the right, the client would tell the server to add one additional column of slices on the right side of the region of interest, expanding the region of interest in order to preserve a guard band of at least one slice all the way around the view window.

The above described embodiment does not take into account the rate at which the user is panning. A more sophisticated embodiment could add additional computational ability to take into account the rate at which the user pans the view window. This added logic could be added at either the server or the client. The following example is based on the logic for rate being at the server. In such a situation the system would operate as follows: Assume that the user starts panning to the right at a rate of 4.5 degrees per frame. The client plug-in would communicate this rate back to the server. Periodically, the client would also communicate back to the server the actual current position of the view window. The server would use this information to predict the probable range of locations the view window may have by the time each frame is actually displayed, and send the slices which cover this range (plus a suitable guard band). Thus, when sending the first "I" frame, the server would send the slices covering the current region of interest and all of the slices anticipated up to where the region of interest will probably be located at the time when the next "I" frame is displayed.

In the above example, this would add two columns of slices to the right, since by the time the next "I" frame is reached, the panning may have progressed through 45 degrees. The first "B" frame following this "I" frame will need to transmit only the same 6 by 18 slice region as transmitted from the "I" frame, since the anticipated motion would not have moved too far. For the next 4 "B" frames, the slices covering the 7 by 18 slice region (adding an additional column to the right) would be sent, and the final 4 "B" frames would include all slices in the 8 by 18 slice region (adding two additional columns to the right). The next "I" frame would need to include a 10 by 18 slice region, in anticipation that it would need to cover the possible motion of the previous "B" frames as well as the future "B" frames. As

1 Another embodiment of the invention is shown in Figure 9: In this embodiment a "Stand
2 Alone Panoramic Video Client" 902 is used. In this embodiment, the function of the server
3 module and the client plug-in are co-located on the same computer. The server component
4 904 called the "Panoramic Media Access Module" retrieves and reads the desired
5 panoramic video from a file system 905 that could be local hard drives, CDs, or a
6 networked file system. This module 904 slices the panoramic video frames in the same
7 way as described in the first embodiment and is functionally equivalent to the module 105 in
8 the first embodiment. The "Panoramic Video Renderer" 903 takes the sliced video frames
9 and renders the image to the screen in the same ways as the plug-in 155 in the first
10 embodiment. The "Sliced Video Stream" is equivalent to that described in the first
11 embodiment. In this case, the stream is passed via an inter-process communications
12 mechanism that could include shared memory, pipes, sockets or an equivalent mechanism
13 instead of being streamed through a public or private network. The "Session Control
14 Stream" is the same as the other embodiments and consists of instructions on how to slice
15 the Video stream as it is read from the file system

16
17 While the invention has been shown and described with respect to preferred embodiments
18 thereof, it should be understood that a wide variety of changes may be made without
19 departing from the present invention. The scope of the invention is limited only by the
20 appended claims: